

Topological Index as Applied to π -Electronic Systems.

II. Topological Bond Order*

Haruo HOSOYA and Miyuki MURAKAMI

Department of Chemistry, Ochanomizu University, Bunkyo-ku, Tokyo 112

(Received June 24, 1975)

Topological bond order p_l^T is proposed to be defined for a bond l in the graph corresponding to the carbon atom skeleton of a hydrocarbon. This quantity is obtained by hand calculation of the topological indices of the graph and its subgraphs. The p_l^T value is linearly correlated well with the Coulson bond order p_l^C for both the chain and cyclic compounds. If the bonds with the same Pauling bond order p_l^P are compared this linear relationship is greatly improved and one gets an empirical relation as

$$p_l^O = A p_l^T + B p_l^P.$$

In order to describe the π -electronic distribution on the carbon atom skeleton of a given unsaturated hydrocarbon several bond orders have been proposed, e.g., by Paluig,¹⁾ Coulson,²⁾ Mulliken,³⁾ Ham and Ruedenberg,⁴⁾ and so on. The Coulson bond order, p_{rs}^C , is obtained from the molecular orbitals of LCAO form as

$$p_{rs}^C = 2 \sum_n^m C_{nr} C_{ns} \quad (1)$$

in terms of the notations currently adopted. Because of mathematical neatness let us confine ourselves to the ground state of neutral unsaturated hydrocarbon molecules with even number $[N=2m]$ of π -electrons without any NBMO and let the term "molecular orbital" be meant by the Hückel molecular orbital. The Pauling bond order, p_l^P , is defined in the valence bond method as

$$p_l^P = \frac{K(G \ominus l)}{K(G)}, \quad (2)$$

where $K(G)$ is the number of the Kekulé structures for graph G (composed of points and lines corresponding to the carbon atom skeleton of a hydrocarbon) and the notation $G \ominus l$ stands for the subgraph of G obtained by deleting a bond, say $l(rs)$, together with all the bonds adjacent to l . The value $K(G \ominus l)$ is equal to the number of the Kekulé structures for G such that a double bond is assigned to l . This means that the Pauling bond order is the weight of a bond in the graph with respect to a set of the Kekulé structures.

Ham has given an ingenious proof⁵⁾ that the molecular orbital (MO) and valence bond (VB) methods are rigorously connected to each other through the bond order p_{rs}^{HR} as defined by

$$p_{rs}^{HR} = 2 \sum_n^m \frac{C_{nr} C_{ns}}{X_n}, \quad (3)$$

where X_n is the eigenvalue of the secular equation, or the root of the characteristic polynomial

$$\begin{aligned} P_G(X) &= (-1)^N \det |A - XE| \\ &= \sum_{k=0}^N a_k X^{N-k}. \end{aligned} \quad (4)$$

Here A and E are respectively the adjacency and unit

matrices of graph G . It was shown that for alternant hydrocarbons without any $4n$ -membered ring the MO expression p_{rs}^{HR} is identical to the VB expression p_l^P ,

$$p_{rs}^{HR} = p_l^P \quad (l = \overline{rs}). \quad (5)$$

Aside from the two main streams (MO and VB) of the electronic theory the topological or graph-theoretical approach is now being realized as a powerful and perspective tool for understanding the nature of the several properties of molecules.^{6,7)} In this paper a topological bond order, p_l^T , is proposed to be defined as

$$p_l^T = \frac{Z_{G \ominus l}}{Z_G}, \quad (6)$$

where Z_G is the topological index for graph G . The topological index is an integer assigned to a graph for characterizing the topological nature of the graph and will be explained later in detail. It will be shown that p_l^T is correlated well with all the bond orders described above. One can roughly estimate the relative magnitudes of the bond orders quite easily with a "paper and pencil" manipulation for the graph.

Topological Index

Consider a graph (G) composed of linearly connected six points, which may be the carbon atom skeleton of n -hexane or 1,3,5-hexatriene (I) (See Table 1). A non-adjacent number $p(G, k)$ is defined as the number of ways in which k disconnected lines are chosen from G , $p(G, 0)$ being defined as unity for all the graphs. It has been shown that for a tree graph, i.e., a graph without any ring, the coefficients of the characteristic polynomial is equal to $p(G, k)$ numbers as

$$P_G(X) = \sum_{k=0}^m (-1)^k p(G, k) X^{N-2k} \quad (G \in \text{Tree}). \quad (7)$$

Especially for linear graph \bar{N} composed of consecutive N points we have

$$P_{\bar{N}}(X) = \sum_{k=0}^m (-1)^k \binom{N-k}{k} X^{N-2k}. \quad (8)$$

These are special cases of a more general relation for non-tree graphs between the $p(G, k)$ and $P_G(X)$,

$$P_G(X) = \sum_{R_i \in G} \sum_{k=0}^{m_i} (-2)^{r_i} (-1)^k p(G \ominus R_i, k) X^{N-n_i-2k} \quad (9)$$

where R_i runs over all the sets of $r_i (\geq 0)$ disjoint rings with a total of $n_i (\geq 0)$ points.^{7,8)} This relation is

* The first paper of this series is, H. Hosoya, K. Hosoi, and I. Gutman, "A Topological Index for the Total π -Electron Energy. Proof of a Generalised Hückel Rule for an Arbitrary Network," *Theor. Chim. Acta*, **38**, 37 (1975).

TABLE 1. PHYSICAL MEANING OF THE TOPOLOGICAL INDEX AS EXEMPLIFIED BY GRAPH 6.

$G \equiv \bar{6}$		$G-1 \equiv \bar{5}$	
k	ℓ	$p(G,k)$	$p(G-1,k)$
0	1 2 3 4 5	1	1
1		5	4
2		6	3
3		1	0
	5 3 4 3 5	13	8
	$Z_{G \ominus \ell}$	Z_G	Z_{G-1}

equivalent to the more compact one derived independently by Sachs

$$P_G(X) = \sum_{n=0}^N \sum_{s \in S_n} (-1)^{c(s)} 2^{r(s)} X^{N-n} \quad (10)$$

where s runs over all the Sachs graph composed of n points, $c(s)$ components and $r(s)$ rings.⁹⁻¹¹ However, Eq. (9) has more advantage over and is more practical than Eq. (10) in that useful recursive relations which will be described later are known to get the $p(G,k)$ numbers for larger graphs.

The value $p(G, 1)$ is nothing else but the number of lines in G . As illustrated in Table 1 for I ($=\bar{6}$) there are six different ways in which two disconnected lines are chosen and we have $p(G, 2)=6$. There is only a unique choice of three disconnected lines for G and one gets $p(G, 3)=1$, the maximum number m of k for I being three. This means graph I has only one Kekulé structure.

The topological index is the sum of the $p(G, k)$ numbers of a given graph as

$$Z_G = \sum_{k=0}^m p(G, k). \quad (11)$$

For tree graphs Z_G is the sum of the absolute values of the coefficients of the characteristic polynomial.

$$\begin{aligned} Z_G &= \sum_{k=0}^N |a_k| \\ &= \sum_{k=0}^m |a_{2k}| \quad (G \in \text{tree}). \end{aligned} \quad (12)$$

For later discussions it is convenient to define the Z-counting polynomial $Q_G(Y)$ as

$$Q_G(Y) = \sum_{k=0}^m p(G, k) Y^k. \quad (13)$$

With this polynomial the topological index is expressed as

$$Z_G = Q_G(1). \quad (14)$$

For graph I in Table 1 we have

$$Q_{\bar{6}}(Y) = 1 + 5Y + 6Y^2 + Y^3 \quad (15)$$

and $Z_{\bar{6}}=13$. Extensive tabulation on the quantities

$p(G, k)$, Z_G , and $P_G(X)$ are given elsewhere.¹²⁻¹⁴

Let such a subgraph of G be denoted as $G-l$ that is derived by deleting bond l . The polynomial $Q_G(Y)$ is expressed as

$$Q_G(Y) = Q_{G-l}(Y) + Y \cdot Q_{G \ominus l}(Y). \quad (16)$$

If graph G is decomposed into fragments L, M, \dots by the deletion of l ,

$$G-l = L \cup M \cup \dots \quad (17)$$

$$L \cap M \cap \dots = 0, \quad (18)$$

the polynomial $Q_{G-l}(Y)$ reduces to their product

$$Q_{G-l}(Y) = Q_L(Y) \cdot Q_M(Y) \cdot \dots \quad (19)$$

and so does $Q_{G \ominus l}(Y)$.

For the topological index we have

$$Z_G = Z_{G-l} + Z_{G \ominus l} \quad (20)$$

$$Z_{G-l} = Z_L \cdot Z_M \cdot \dots \quad (21)$$

Note that Eqs. (16) and (20) hold true irrespective of the choice of l .

As a special case the Z-counting polynomials $Q_{G-}(Y)$'s of a series of graphs \bar{N} recurse as

$$Q_{\bar{N}}(Y) = Q_{\bar{N}-1}(Y) + Y \cdot Q_{\bar{N}-2}(Y) \quad (N \geq 2) \quad (22)$$

$$Q_{\bar{1}}(Y) = Q_{\bar{0}}(Y) = 1, \quad (23)$$

and their topological indices $Z_{\bar{N}}$'s form the well-known Fibonacci numbers (1,1,2,3,5,8,13...) as

$$Z_{\bar{N}} = Z_{\bar{N}-1} + Z_{\bar{N}-2} \quad (N \geq 2) \quad (24)$$

$$Z_{\bar{0}} = Z_{\bar{1}} = 1. \quad (25)$$

A number of interesting mathematical relations are derived for $Q_G(Y)$ and Z_G .^{6,15}

Topological Bond Order

The value Z_G may be regarded as the total number of ways in which disconnected double bonds are assigned onto the single bond skeleton of G as in Table 1, where a 13×5 matrix is constructed with elements o and x corresponding respectively to double and single bonds. Each row forms a set of disconnected double bonds and the number of rows gives the topological index Z_G . Then this number is thought to represent the degree of choice for distributing π -electrons over the molecular skeleton G . The mode of π -electron population will then be dependent on the topology of molecules.

Choose any bond l from G , say bond 1 from this graph. Then the Z_G sets of disconnected double bonds are divided into two groups, according to such a criterion that bond l is chosen as double (l -inclusive) or not (l -exclusive). The numbers of these two groups are denoted respectively as Z_G (l -inclusive) and Z_G (l -exclusive) and we have

$$Z_G = Z_G(l\text{-inclusive}) + Z_G(l\text{-exclusive}). \quad (26)$$

They are obtained respectively as the numbers of o and x's in the given column l . In this case ($l=1$) one gets $Z_G(l\text{-inclusive})=5$ and $Z_G(l\text{-exclusive})=8$. It is to be noted that each set of disconnected double bonds in the latter group appears once and only once in counting the value of Z_{G-l} (See the double rectangles in

TABLE 2. COMPARISON OF THE TOPOLOGICAL AND COULSON BOND ORDERS FOR AN ISOMER II OF TETRADECAHEPTAENE^{a)}

Bond type	Bond ^{a)} l	$Z_{G\ominus l}^{b)}$	p_l^c	
			approx. ^{c)}	exact
Single bond	f	70	0.387	0.389
	b	75	0.400	0.400
	h	81	0.416	0.411
	d	100	0.466	0.469
	j	105	0.479	0.478
	l	109	0.489	0.488
Double bond	e	125	0.713	0.709
	k	148	0.775	0.777
	c	150	0.780	0.780
	g	162	0.813	0.818
	m	183	0.869	0.867
	i	197	0.907	0.907
	a	200	0.915	0.913

a) See Fig. 1. b) $p_l^T = Z_{G\ominus l}/Z_G$, $Z_G = 475$ for this graph.

c) For single bonds: $0.002636 Z_{G\ominus l} + 0.2020$, for double bonds: $0.002697 Z_{G\ominus l} + 0.3757$.

Table 1). It means

$$Z_{G-l} = Z_G(l\text{-exclusive}). \quad (27)$$

Now by combining Eqs. (20), (26), and (27) one gets

$$\begin{aligned} Z_{G-l}(\text{inclusive}) &= Z_G - Z_{G-l} \\ &= Z_{G\ominus l}. \end{aligned} \quad (28)$$

The quantity $Z_G(l\text{-inclusive})$, or $Z_{G\ominus l}$, can then be considered as the weight of bond l in distributing π -electrons over G , and the topological bond order for l in G may be defined as Eq. (6).

In what follows a number of examples will be given of the good correlation between p_l^c and p_l^T or $Z_{G\ominus l}$. In Table 2 and Fig. 1 the values of p_l^c and $Z_{G\ominus l}$ for an unsaturated hydrocarbon isomer II with 14 electrons

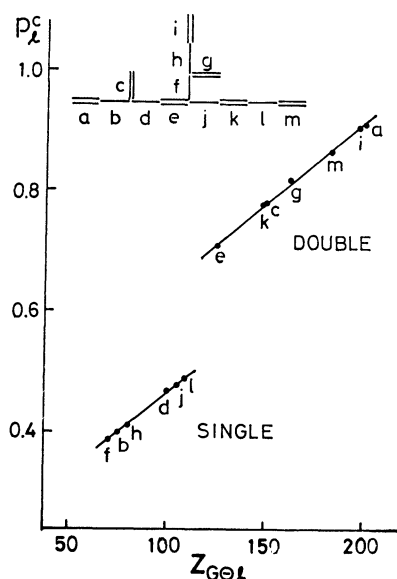


Fig. 1. Linear relationships between the Coulson bond order p_l^c and the topological bond order p_l^T . $p_l^T = Z_{G\ominus l}/Z_G$. The Z_G value for this graph (II) is 475. See also Table 2.

are compared. Note that two different sets of linear relationships are observed respectively for the bonds marked with "single" and "double" in the Kekulé structure. A procedure for obtaining the $Z_{G\ominus l}$ values with the recursive relations (20) and (21) is exempli-

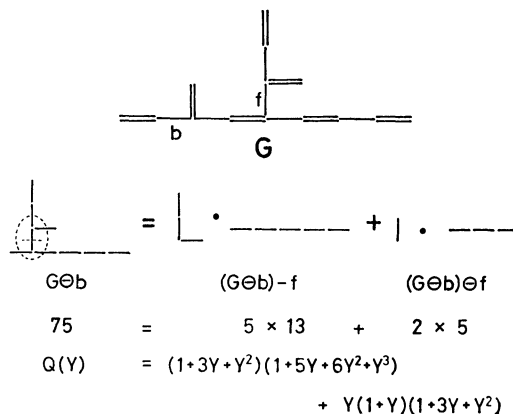


Fig. 2. The procedure for obtaining the topological index Z_G and the Z -counting polynomial $Q_G(Y)$ as exemplified by the subgraph $G\ominus b$ of the graph given in Fig. 1. Bond f is chosen to decompose the graph into fragments.

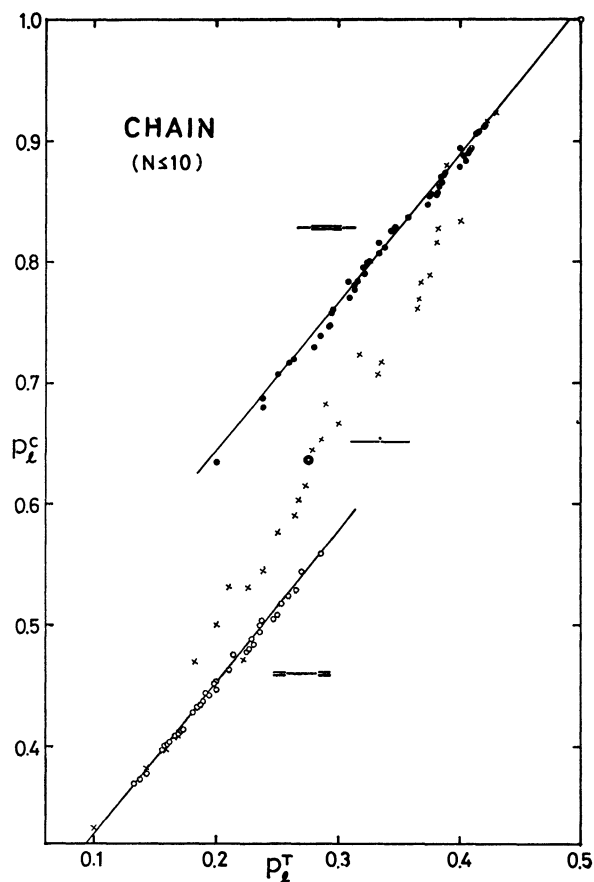


Fig. 3. Plots of p_l^c and p_l^T values for all the possible isomers of polyenes not larger than decapentaene. ○: single bond, ●: double bond, x: bond in a radical, ⊙: limiting bond order ($p_l^c = 0.6366$ and $p_l^T = 0.2764$) of an infinite polyene (See Eqs. (29)–(31), (40), (43)).

fied in Fig. 2 for bond *b* in the graph concerned, where bond *f* is chosen to decompose the graph into fragments. The topological indices of the fragmental graphs are given as the Fibonacci numbers or their combinations. The procedure is simple and easy enough for hand calculation but the result is surprisingly predictable.

Figure 3 gives the plots of p_i^c and p_i^r for all the possible isomers of chain polyenes with not more than ten carbon atoms. The circles for the "single" (open) and "double" (filled) bonds in even-numbered polyenes with singlet ground states lie respectively on the two parallel lines, while for radicals (marked with crosses) the points lie on a line spanning from the "single bond line" over to the "double bond line" in Fig. 3. Two different classes of radicals are involved in this category, *i.e.*, even-numbered polyenes with a triplet ground state as III and odd-numbered polyenes with a doublet or higher multiplet ground state as IV. Thus for



chain polyenes or hydrocarbon radicals the Coulson bond order p_i^c can be estimated fairly accurately by hand calculation of p_i^r value and bond type (single, double, or radical) classification. Tentatively from the set of points in Fig. 3 the following relations were obtained for the three classes of bonds.

$$p_i^c = 1.24 p_i^r + 0.392 \quad (l=\text{double bond}) \quad (29)$$

$$p_i^c = 1.24 p_i^r + 0.200 \quad (l=\text{single bond}) \quad (30)$$

$$p_i^c = 1.67 p_i^r + 0.164 \quad (l=\text{bond in a radical}). \quad (31)$$

It is interesting to note that the topological bond orders of a linear polyene oscillate by going from the terminal to the center and converge to a limit as predicted by the p_i^c value. This is demonstrated in Fig. 4 for 1,3,5,7,9,11,13-tetradecaheptaene. For an infinite polyene p_i^r and p_i^c values converge to $(5-\sqrt{5})/10=0.2764$ and $2/\pi=0.6366$ respectively (See Appendix), the speed of convergence of the former being much faster. Thus it is expected that for higher members of chain polyenes the relations (29)–(31) are getting worse as shown in Fig. 5, where p_i^r and p_i^c values for

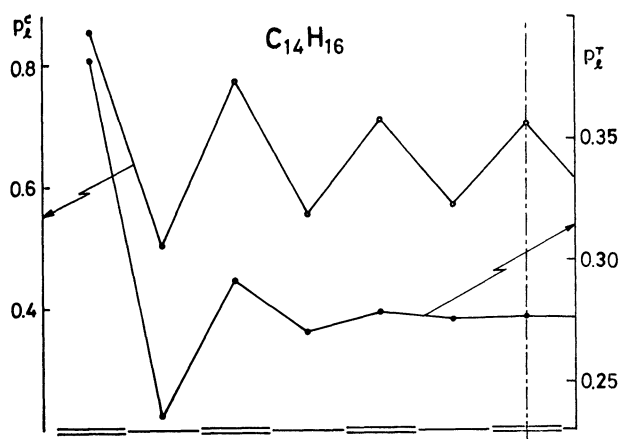


Fig. 4. Zig-zag nature of the Coulson and topological bond orders for a linear polyene (1,3,5,7,9,11,13-tetradecaheptaene).

linear polyenes larger than decapentaene are plotted. This is the limitation of the topological bond order for chain compounds.

For cyclic compounds there is no rigorous one-to-one correspondence between the coefficients of $P_G(X)$ (Eq. (7)) and $Q_G(Y)$ (Eq. (13)) polynomials as inferred from Eq. (9). Then correlation between p_i^c and p_i^r values is not expected as good as what is

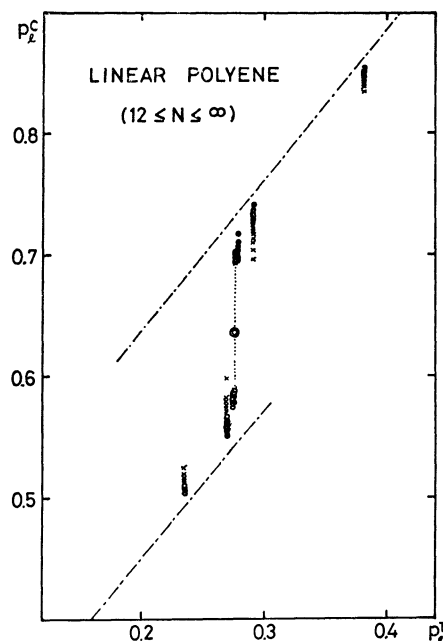


Fig. 5. Plots of p_i^c and p_i^r values for larger linear polyenes ($N > 10$). In the dotted line region a number of circles and crosses are crowded. The two parallel lines are the "single" and "double bond lines" in Fig. 3. Other marks are the same as used in Fig. 3.

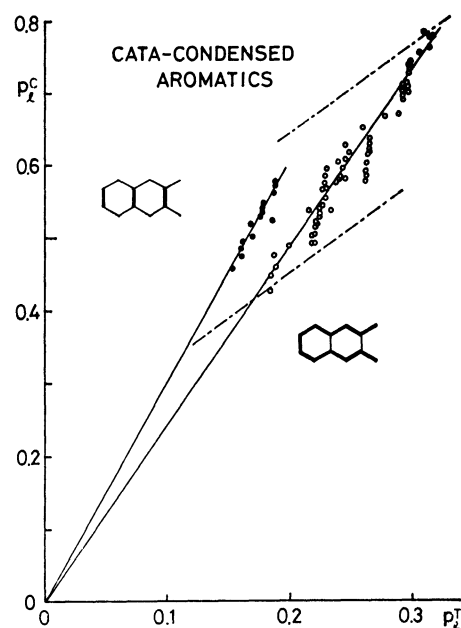


Fig. 6. Plots of p_i^c and p_i^r values for cata-condensed aromatic hydrocarbons not larger than pentacene. ●: bridge bond, ○: peri-bond,

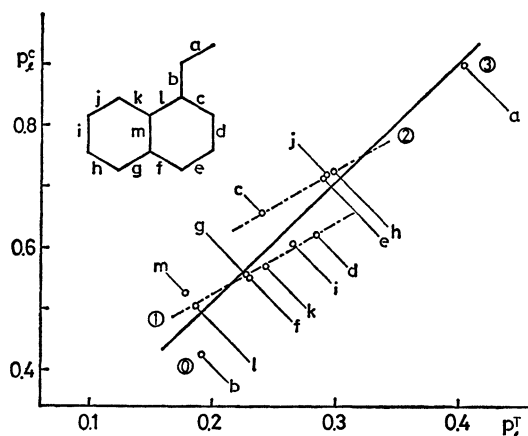


Fig. 7. Correlation between the p_i^C and p_i^P values of 1-vinylnaphthalene. The number in the circle indicates the $K(G \ominus l)$

shown in Figs. 1 and 3. In Fig. 6 are plotted the p_i^C and p_i^P values for cata-condensed aromatic hydrocarbons not larger than pentacene. Almost all the points lie within the region between the two parallel lines for single and double bonds in Fig. 3 and are grouped into two parts. The filled and open circles respectively correspond to the "bridge" and "peri-bonds". Both the classes of points lie on the respective lines passing through the origin. The linearity of these two classes of points is not so good as that of small polyenes (Fig. 3) but is unexpectedly good if one thinks of a simple procedure for getting the p_i^P values. Further, on the contrary to the chain compounds, the larger the cyclic graphs the better the correlation between p_i^C and p_i^P values as shown in Fig. 7 for 1-vinylnaphthalene. The thirteen points lie around a straight bold line. Inspection of this figure reveals an interesting feature of the correlation between p_i^C and p_i^P values. That is, the "corner bonds" (c, e, h, and j) lie perfectly on a straight line, while the "side bonds" (d, f, g, i, k, and l) lie on another line parallel to the former. The point for the "bridge bond" m is located nearby the latter line. Note that the unnormalized Pauling bond orders p_i^P 's, i.e., $K(G \ominus l)$ defined in Eq. (2), of these two groups of bonds are respectively two and one. The $K(G \ominus l)$ values of bonds a and b are respectively three and zero. Similar relations have been found in Eqs. (29)–(31).

Now it is empirically induced that for a certain group of molecules or for a group of bonds in a certain molecule the following relation is generally observed,

$$p_i^C = A p_i^P + B p_i^T. \quad (32)$$

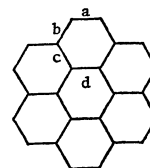
This will be mathematically proved in a forthcoming paper by the use of an integration over the complex plane.^{7,16)}

Recently it is discussed which of the bond orders p_i^C and p_i^P is correlated better with the bond length.^{17,18)} However, as the aim of the present paper is to show the relation between these bond orders, much will not be discussed on this problem. Table 3 gives the comparison of the bond lengths and various bond orders for coronene. This is one of the cases where the topological bond order shows the best correlation

TABLE 3. COMPARISON OF THE BOND LENGTHS AND VARIOUS BOND ORDERS FOR CORONENE

Bond	Coulson p_i^C	Pauling p_i^P	Topological p_i^T	Bond length ^{a)} (Å)
a	0.745	0.7	0.314	1.347
b	0.538	0.3	0.222	1.417
c	0.538	0.4	0.185	1.433
d	0.522	0.3	0.206	1.425

a) J. M. Robertson and J. G. White, *Nature*, **154**, 605 (1944); J. K. Fawcett and J. Trotter, *Proc. Roy. Soc., Ser. A*, **289**, 366 (1966).



with the experimentally obtained bond lengths.

Appendix

The Limiting Values of the Topological Bond Order p_i^T .

Consider a polyene \bar{N} with N consecutive points linearly connected. The topological bond order p_i^T of the l -th bond from one of the terminals is given by Eq. (6) as

$$p_i^T = \frac{Z_{\bar{N} \ominus l}}{Z_{\bar{N}}}. \quad (33)$$

Note that graph $\bar{N} \ominus l$ has two component graphs $\bar{l}-1$ and $\bar{N}-l-1$.

Then we have

$$Z_{\bar{N} \ominus l} = Z_{\bar{l}-1} \times Z_{\bar{N}-l-1}. \quad (34)$$

The $Z_{\bar{N}}$ series form the Fibonacci numbers and is expressed by the Binét's formula as

$$Z_N = (\alpha^{N+1} - \beta^{N+1})/\sqrt{5} \quad (35)$$

where

$$\alpha = (1 + \sqrt{5})/2$$

and

$$\beta = (1 - \sqrt{5})/2. \quad (36)$$

From Eqs. (33)–(36) one gets

$$p_i^T = \frac{1}{\sqrt{5} \alpha} \frac{1 - X^l + X^N(1 - X^{-l})}{1 - X^{N+1}} \quad (37)$$

where

$$X = \beta/\alpha = (\sqrt{5} - 3)/2 = -0.38197. \quad (38)$$

For larger N Eq. (37) reduces to

$$p_i^T \rightarrow (1 - X^l)/(\sqrt{5} \alpha) \quad (39)$$

which oscillates by going from the terminal ($l=1$) to the center ($l \rightarrow \infty$) and finally reaches

$$p_i^T \rightarrow 1/(\sqrt{5} \alpha) = (5 - \sqrt{5})/10 = 0.27639. \quad (40)$$

This behavior of p_i^T corresponds to that of p_i^C as expressed by

$$p_i^C = \frac{1}{N+1} \left\{ \operatorname{cosec} \frac{\pi}{2N+2} + (-1)^{l-1} \operatorname{cosec} \frac{(2l+1)\pi}{2N+2} \right\} \quad (N=\text{even}) \quad (41)$$

$$p_i^c = \frac{1}{N+1} \left\{ \cot \frac{\pi}{2N+2} + (-1)^{l-1} \cot \frac{(2l+1)\pi}{2N+2} \right\} \quad (N=\text{odd}).^{19)} \quad (42)$$

Namely, since for larger N $\text{cosec } \theta$ and $\cot \theta$ converge to θ^{-1} , one gets

$$p_i^c \rightarrow \frac{2}{\pi} \left\{ 1 + (-1)^{l-1} \frac{1}{2l+1} \right\} \rightarrow 2/\pi = 0.63662. \quad (43)$$

References

- 1) L. Pauling, "The Nature of the Chemical Bond," Cornell Univ. Press, Ithaca, N. Y. (1960), p. 232.
- 2) C. A. Coulson, *Proc. Roy. Soc., Ser. A*, **169**, 413 (1939).
- 3) R. S. Mulliken, *J. Chem. Phys.*, **23**, 1833 (1955).
- 4) N. S. Ham and K. Ruedenberg, *J. Chem. Phys.*, **29**, 1215 (1958).
- 5) N. S. Ham, *J. Chem. Phys.*, **29**, 1229 (1958).
- 6) H. Hosoya, This Bulletin, **44**, 2332 (1971).
- 7) H. Hosoya, K. Hosoi, and I. Gutman, *Theor. Chim. Acta*, **38**, 37 (1975).
- 8) H. Hosoya, *Theor. Chim. Acta*, **25**, 215 (1972).
- 9) H. Sachs, *Publ. Math. Debrecen*, **11**, 199 (1963).
- 10) A. Graovac, I. Gutman, N. Trinajstić, and T. Živković, *Theor. Chim. Acta*, **26**, 67 (1972).
- 11) R. B. Mallion, N. Trinajstić, and A. J. Schwenk, *Z. Naturforsch.*, **29a**, 1481 (1974).
- 12) K. Mizutani, K. Kawasaki, and H. Hosoya, *Natural Science Report of Ochanomizu Univ.*, **22**, 39 (1971).
- 13) K. Kawasaki, K. Mizutani, and H. Hosoya, *Natural Science Report of Ochanomizu Univ.*, **22**, 181 (1971).
- 14) T. Yamaguchi, M. Suzuki, and H. Hosoya, *Natural Science Report of Ochanomizu Univ.*, **26**, 39 (1975).
- 15) H. Hosoya, *Fibonacci Quarterly*, **11**, 255 (1973).
- 16) C. A. Coulson and H. C. Longuet-Higgins, *Proc. Roy. Soc., Ser. A*, **191**, 39 (1947).
- 17) N. L. Allinger and J. T. Sprague, *J. Amer. Chem. Soc.*, **95**, 3893 (1973).
- 18) W. C. Herndon, *ibid.*, **96**, 7605 (1974).
- 19) See for example, C. A. Coulson and A. Streitwieser Jr., "Dictionary of π -Electron Calculations," Pergamon Press, Oxford (1965).